WORKING PAPER SERIES

A Reconsideration of the Properties of the Generalized Method of Moments in Asset Pricing Models.

Christopher J. Neely

Working Paper 1994-010A
http://research.stlouisfed.org/wp/1994/94-010A.pdf

Revised Version: 95-018A

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
411 Locust Street
St. Louis, MO 63102

A RECONSIDERATION OF THE PROPERTIES OF THE GENERALIZED METHOD OF MOMENTS IN ASSET PRICING MODELS

April 1994

ABSTRACT

This paper tests the small sample properties of Hansen's (1982) Generalized Method of Moments (GMM) on simulated data from a consumption based asset pricing model. In finite samples the estimates of the coefficient of relative risk aversion and the discount parameter are strongly biased due to the unusual shape of the GMM criterion function for the model and the GMM test statistics perform poorly. In fact, the finite sample properties of the test statistics suggest the rejection results achieved by applying GMM to representative agent asset pricing models with real data (Hansen and Singleton 1982) must be viewed with some circumspection.

KEYWORDS: Generalized method of moments, Monte Carlo simulation Markov Chain

JEL CLASSIFICATION: C52, E44, E47, G12

Christopher J. Neely
Economist
Federal Reserve Bank of St. Louis
411 Locust Street
St. Louis, MO 63102

1. INTRODUCTION

Much recent research in asset pricing has centered on the representative agent framework developed by Lucas (1978). Use of the Lucas framework permits the study of the relationship between movements in output and equilibrium asset prices in a one-good, pure exchange economy.

There have been two main approaches to the use of representative agent models in the study of asset prices. The first approach is known as calibration and was popularized by Mehra and Prescott (1985). The second approach, exemplified by Hansen and Singleton (1982), is to estimate the model directly from the data and conduct formal econometric tests of overidentifying restrictions

Calibration exercises generate simulated data that have properties exhibited by data from the real world. Mehra and Prescott (1985), for instance, calibrated the model in an unsuccessful attempt to produce the high equity premium. More recently, Kandel and Stambaugh (1990) describe a model economy in which the distribution of consumption growth is lognormal. The mean and variance of the distribution depend on the state of nature which is determined by the outcome of a Markov chain process. The state of nature and consumption determine asset prices in this economy. Kandel and Stambaugh (1990) find the model generates simulated data which is consistent in important ways with data from the real economy. Specifically, the conditional mean and variance of consumption growth covary with the business cycle (as in the real economy) and they investigate the implications of this covariance for conditional moments of asset returns. By incorporating the asymmetry of the business cycle and a high coefficient of relative risk aversion in their data generation process they are able to reproduce the pattern of autocorrelations over the return horizon seen in real data. Likewise, Cecchetti, Lam and Mark (1990) construct a similar model and conclude that the negative serial correlation in the data could plausibly have come from such a model.

The second approach is typified by Hansen and Singleton's (1982 and 1984) work describing the Generalized Method of Moments (GMM) for the estimation and testing of models using orthogonality conditions implied by stochastic Euler equations. They apply the GMM to the representative agent model using various monthly measures of consumption growth and asset returns data from the U.S. from 1959:2 to 1978:12 and strongly reject the overidentifying restrictions on the model. This indicates the model is not a good description of the data. The main advantages of GMM are its flexibility and generality in not requiring that the stochastic equilibrium be completely specified. For example, the distributional structure of the error terms does not have to be specified. Due to this generality, the GMM does not require a joint hypothesis about the nature of the underlying economy and the stochastic environment and therefore is the most problematic of the rejections for the use of the representative agent framework.

The GMM rejections from formal tests of the model using real data would seem to make the data coming from simulated representative agent economies much less relevant. However, the size of the formal tests derived from estimators such as the GMM estimator are based on asymptotic results and may be strongly prone to overrejection in small samples from these environments.

Work on the small sample properties of the GMM in the Mehra and Prescott (1985) asset pricing framework has been done by Tauchen (1986) and Kocherlakota (1990). Investigating the properties of two-stage GMM, Tauchen found that (in a specific environment) use of shorter lags in the instrument set produced nearly asymptotically optimal parameter estimates and that the test of overidentifying restrictions performed well in small samples. Kocherlakota (1990) chose to investigate multi-stage GMM because of evidence that the small sample performance of the multi-stage estimators were superior to those of the two-stage estimators. Using different sets of parameter values, Kocherlakota (1990, p. 285) found that "assuming the large sample properties of ... GMM estimators are true can lead one to 'overreject' the model." The properties of GMM in the more elaborate framework of Kandel and Stambaugh (1990) - where the distribution of consumption growth is governed by the state of nature which evolves according to a Markov process - may be quite different from those in the Mehra and Prescott economy and lead to much different conclusions about the use of GMM with representative agent models.

Investigation of the properties of GMM in this environment is worthwhile because Kandel and Stambaugh (1990) and Cecchetti, Lam and Mark (1990) have created models which approximate reality in useful ways. Both papers have shown that observed patterns of autocorrelation in asset returns can be produced from simple representative agent asset pricing frameworks. While it has been known that such behavior does not necessarily imply a violation of market efficiency, it is illustrative to see simple rational expectations models which produce such consequences.

There are two potential goals for this paper. The first goal is to study the small sample properties of GMM as an econometric test in a specific environment. The second goal is to study properties of the representative agent model, specifically whether the representative model produces data that resembles real world data in producing rejections in formal (GMM) tests of the restrictions implied by the model. This paper will focus on the first objective.

This paper extends previous studies of the GMM by Tauchen (1986) and Kocherlakota (1990) in three important ways. First, the representative agent framework of Kandel and Stambaugh (1990) differs from those frameworks studied previously in that consumption growth has a *continuous* (rather than discrete) distribution. Second, the model economy studied here is calibrated according to the Kandel and Stambaugh framework which uses a much higher coefficient of relative risk aversion (CRRA = 55!) than previous works. The changes in the framework are found to substantially alter the properties of the GMM. Finally, the properties of both two-stage and multi-stage GMM are investigated in this environment.

The rest of the paper is organized as follows: Section 2 describes the representative agent framework of Kandel and Stambaugh (1990). Hansen's (1982) GMM technique is explained in Section 3. Section 4 summarizes the results of GMM estimation of the model applied to data from the real world. Section 5 examines the properties of GMM in the simulated environment. Conclusions are drawn in Section 6.

2. THE SIMULATED ECONOMY

2.1 The Utility Function

The Kandel and Stambaugh's (1990) model is based on that of Lucas (1978) who posited an infinitely lived, representative agent who maximizes expected time-additive utility. The agent chooses consumption and a portfolio of N assets in each period to solve the following problem:

$$\max_{c_\tau} E\left(\sum_{\tau=t}^{\infty} \beta^{\tau-t} \cdot \frac{c_\tau^{(1-\alpha)} - 1}{(1-\alpha)} \right), \tag{1}$$

where $0 < \beta < 1$ and $0 < \alpha < \infty$. The agent is subject to the following budget constraint at time *t*:

$$c_t + p_t' q_t \leq p_t' q_{t-1} + w_t \tag{2}$$

where p_t, d_t and q_t are N by 1 vectors of asset prices, dividends paid and asset quantities held at time *t*. w_t is labor income at time *t*. Maximization of (1) subject to (2) produces the following first order necessary condition for each asset

$$1 = \beta \cdot E_t \left[\left(\frac{p_{t+1} + d_{t+1}}{p_t}\right) \cdot \left(\frac{c_{t+1}}{c_t}\right)^{-\alpha} \right] \tag{3}$$

The utility function U(c) is of the constant relative risk aversion class. The parameter α is the coefficient of relative risk aversion. This parameter measures the curvature of the utility function, the agent's tolerance for risk and desire to smooth consumption intertemporally. The standard assumption of constant relative risk aversion ensures the equilibrium return process is stationary.

2.2 Consumption Growth

Aggregate consumption in each period is equal to aggregate output each period. Consumption growth in period t is denoted by λ_t. The innovation in Kandel and Stambaugh's version of this model is that the distribution of consumption growth is continuous. The parameters depend on the state of nature which is governed by a Markov process: $\ln(\lambda_t)$ is normally distributed with mean μ_t and variance σ^2_t. The parameters of the distribution of $\ln(\lambda_t)$ (μ_t and σ_t) are a function of the state of nature at time t. The number of discrete states of nature is denoted by s. The state of nature evolves according to a finite dimension, ergodic Markov chain. The transition matrix for the state of nature is denoted by phi (ϕ). The transition matrix and the parameters of the consumption growth process are given in the Appendix.

In each period, the state of nature is chosen by a markov process, then the realization of consumption growth for that period is drawn from the appropriate distribution (dependent on the realized value for the state). The state of nature and consumption growth determine asset prices (hence asset returns) for that period.

2.3 The Asset Prices

The asset prices in each period are given by functions of the state of nature and consumption growth. The infinite future discounted stream of asset dividends depend only on the current state and current consumption. Therefore, a sufficient statistic for the state of the economy in any period is the pair (c,i), where c is the consumption in the period and i indexes the state of nature.

Kandel and Stambaugh consider three types of assets, a risk free bond, a share of aggregate wealth, and a share of levered equity. Levered equity is a share of aggregate wealth minus a claim on a risky bond. To derive closed form solutions for asset returns, one uses the property that the price of an asset is equal to the present value of the expected discounted future dividend stream and the Euler equation (3). The solutions to the asset pricing equations presented below are discussed extensively in Mehra and Prescott (1985) and Kandel and Stambaugh (1988). They will be presented below with minimal explanation.

The payoff on the risk free asset is one unit of the consumption good. Substituting this in equation (3) for $p_{t+1} + d_{t+1}$ and using the property of the Markov transition matrix for the state of nature. It may be shown that the gross return on the risk free asset when the economy goes from state i to state j is given by

$$R^f = \frac{1}{\beta \cdot \sum_{j=1}^{s} \phi_{ij} \cdot E[\lambda(i)^{-\alpha}]} = \frac{1}{\beta \cdot E[\lambda(i)^{-\alpha})]} \tag{4}$$

where ϕ_{ij} denotes the ith row, jth column of the Markov transition matrix for the states. λ(i) is consumption growth if the state of nature is *i*.

The payoff to one share of aggregate wealth is, of course, a claim to all consumption in perpetuity. Using the same sort of reasoning used to derive the risk free return above, the return on aggregate wealth (going from state i to state j) is given by

$$R^{aw} = \lambda(i) \cdot \frac{(1 + w_j)}{w_i} \tag{5}$$

where w_i is the ith element of the s vector w given by

$$w = (I - H)^{-1} \cdot H \cdot 1_s \tag{6}$$

In equation (6), H is an $s \times s$ matrix with (i,j)th element

$$h_{ij} = \beta \cdot \phi_{ij} \cdot E[\lambda(i)^{1-\alpha}] \tag{7}$$

and 1_s is an $s \times 1$ vector of ones.

Levered equity is an asset consisting of a share of aggregate wealth net of a risky bond. That is, the holder of a share of levered equity has sold a bond to purchase a share of aggregate wealth and will pay off according to the realization of the return to aggregate wealth. The payoff on a risky bond bought at time t is a fraction (θ) of aggregate wealth at time $t+1$ if aggregate wealth at time $t+1$ is greater than or equal to θ times the value of aggregate wealth at time t. If aggregate wealth at time $t+1$ is less than θ times the value of aggregate wealth at time t, then the risky bond holder receives all wealth at time t. The return on levered equity is a complicated nonlinear function due to the nature of its risky payoff. Before defining the return on levered equity, first define the $s \times 1$ vector

$$g = Y \cdot 1_s \tag{8}$$

where the Y is the $s \times s$ matrix with (i,j)th element

$$y_{ij} = \beta \cdot \phi_{ij} \cdot E[\min[\lambda(i)^{1-\alpha} \cdot (1 + w_j),\ \lambda(i)^{-\alpha} \cdot \theta \cdot w_i]\ ] \tag{9}$$

Then the gross return on levered equity may be given by

$$R^{le} = \frac{\max[0\ ,\ \lambda(i) \cdot (1 + w_j) - \theta \cdot w_i\ ]}{(w_i - g_i)} \tag{10}$$

2.4 Calibrating the Model

Kandel and Stambaugh chose the parameters of their model economy in order to match the first two moments of consumption growth and the value-weighted New York Stock Exchange returns and the expected value of T-bill returns from quarterly data from the real economy. They chose parameter values of $\beta = .9973$, $\alpha = 55$, and $\lambda = .478$. Kandel and Stambaugh argue the parameters should be chosen to match the data rather than ex ante expectations about the correct values of the parameters.

The value of $\alpha = 55$ seems extraordinarily high at first glance, but it is necessary to produce the desired equity premium and interest rate. Objections to such a high value of α are usually predicated on the results of thought experiments (Kandel and Stambaugh 1988). For example, a value of $\alpha = 55$

means that a person with an income of $50,000 would pay $9,483 to avoid an even bet of $10,000. But the problem is with these experiments is the assumption of *constant* relative risk aversion. Given appropriate sizes of a bet, almost any level of risk aversion could seem plausible or implausible. A value of $\alpha = 2$ is usually considered reasonable, but it means that a person with the same wealth would pay only $1.25 to avoid an even bet of $250. A person with $\alpha = 55$ would pay a more plausible $33.96 to avoid the same bet.

2.5 Results Using this Framework

Kandel and Stambaugh simulated data sets of 211 observations (approximately 53 years of quarterly data) and compared the behavior of their simulated data to real data. Their simulated data exhibited the skewness and kurtosis typical of real consumption growth and asset return data. In addition, they found that they were able to reproduce the "U" shaped pattern of autocorrelation of equity returns over return horizons using the equilibrium model of rational behavior described above. That is, they produced data whose returns exhibited low negative first order autocorrelation for returns at short horizons, more negative first order autocorrelations at longer horizons and less negative first order autocorrelations for returns at longer horizons.

Using a similar model of the economy, Cecchetti, Lam and Mark (1990) point out that although "It is well known that serial correlation of returns does not in and of itself violate market efficiency. Nevertheless, there is a tendency to conclude that evidence of mean reversion in stock prices constitutes a rejection of equilibrium models of rational asset pricing." The rational asset pricing model constructed by Cecchetti, Lam and Mark (1990) produces data whose returns are negatively serially correlated. This illustrates "that negative serial correlation in long horizon stock returns is consistent with an equilibrium model of asset pricing."

3. THE GENERALIZED METHOD OF MOMENTS

The use of GMM to test models and estimate parameters is described by Hansen and Singleton (1982). They use the method to test and reject the model using actual monthly consumption and returns data. More on their results is provided in section 5.

3.1 Orthogonality Conditions

Intuitively, the idea behind GMM is to take orthogonality conditions implied by the model (such as those implied by equation (3)) and find parameter values of the model which make sample counterparts close to zero according to some optimal metric. A quadratic function of the orthogonality conditions is constructed and minimized to find suitable values of the parameters.

The following discussion relies heavily on Hansen (1982). We assume that the econometrician has T+1 time series observations on N gross returns, K instruments of some interest, and consumption growth. Let $\pi = \{ \alpha, \beta \}$ be the vector of parameters of interest and π^* the true values of those parameters. Also, denote the vectors of consumption growth by λ, gross returns by R, and the matrix of instruments by Z. In practice the instruments used are generally lagged values of R, λ and a constant. Finally, let the function $f(R_t, \lambda_t, \pi)$ be a function such as that implied by the Euler equations (3)

$$E_{t-1}\left[f(R_t, \lambda_t, \pi^*)\right] = 0 \tag{11}$$

for all t. In the present case, for example

$$f(R_t, \lambda_t, \pi) = [\, \beta \cdot \lambda_t^{-\alpha} * R_t - 1] * Z_t \tag{12}$$

where the symbol * denotes element by element multiplication. Define the function g_T to be the mean of the sample orthogonality conditions.

$$g_T(\pi) = \frac{1}{T} \cdot \sum_{t=1}^{T} f(R_t, \lambda_t, \pi) \tag{13}$$

Hansen shows that under the null g_T converges almost surely to zero as T goes to infinity.

3.2 Weighting Matrix

Hansen shows that for any arbitrary (NK by NK) weighting matrix W of full rank

$$\hat{\pi} = argmin_{\pi}\ g_T(\pi)' \cdot W \cdot g_T(\pi) \tag{14}$$

is a consistent, asymptotically normal estimator of π^*, the true parameters. Further, Hansen shows that the weighting matrix W given by

$$W = W^* = [\ E\ (\ f(R,\lambda,\pi) \cdot f(R,\lambda,\pi)'\ )\ ]^{-1} \tag{15}$$

produces $\hat{\pi}$ that is asymptotically efficient among the class of estimators produced by linear combinations of the orthogonality conditions. Note that the optimal weighting matrix defined by (15) is the inverse of the variance-covariance matrix of the NK orthogonality conditions.

We can produce a consistent estimator of W^*, W_T, by using a consistent estimator $\hat{\pi}$ of π^* in the sample analogue of (14) with an identity matrix used as W. Hence, given a consistent estimator $\hat{\pi}$ of π^* we can get an estimator of π^* that is both consistent and asymptotically efficient in its class. The estimator is

$$\hat{\pi}^* = argmin\ g_T(\pi)' \cdot W_T(\hat{\pi}) \cdot g_T(\pi) \tag{16}$$

The above estimator is known as the two-stage GMM estimator.

By repeatedly iterating over (14), (15) and (16) until the weighting matrix stops changing, we implement multi-stage GMM. Both two-stage and multi-stage GMM estimation are used in this paper.

3.3 Test of Overidentifying Restrictions

Hansen shows further that under the null hypothesis that the model is true that the J statistic given by

$$J_T = T \cdot [g_T(\hat{\pi}^*)' \cdot W_T(\hat{\pi}^*) \cdot g_T(\hat{\pi}^*)] \tag{17}$$

has an asymptotic chi square distribution with NK - 2 degrees of freedom. This statistic is used to test the overidentifying restrictions of the model implied by the orthogonality conditions. The intuition behind the J statistic is that if the model fits the data well, the sample counterparts of the orthogonality conditions can be made close to zero, the J statistic will be small and we will be unable to reject the null hypothesis that the model is correct.

4. RESULTS USING DATA FROM THE REAL WORLD

Hansen and Singleton (1982) test the overidentifying restrictions implied by the representative agent asset pricing model using monthly consumption and returns data. For many of the combinations of instrument sets and asset returns they studied, they were able to reject the null hypothesis that the model was true. Because the GMM does not require a complete specification of the economy, this rejected the Kandel/Stambaugh model of the data generating process.

In order to facilitate the comparison to the Kandel/Stambaugh model economy which is calibrated for quarterly data, the null hypothesis was retested using two-stage and multi-stage GMM on quarterly data. The data were constructed from T-bill, consumption (nondurables and services) and population data taken from Federal Reserve data bases. Value-weighted New York stock exchange data were obtained from the CRSP tapes. Nominal returns were converted to real returns using the implicit consumption deflator. The data ran from 1959:1 to 1989:4, 120 quarterly observations.

Table 1 describes the 6 combinations of instrument sets and returns used on the quarterly data from the real world. The results from these combinations are shown in Table 2. The initial values for the coefficients were those of the hypothesized model economy, $\beta = 99731$, $\alpha = 55$.

As discussed in Tauchen (1986) and Kocherlakota (1990) the point estimates of α and β are quite sensitive to the choice of asset and/or instrument sets. The J statistics frequently reject the model particularly for those combinations that included T-bill returns. These results are broadly consistent with those of Hansen and Singleton (1982) who applied GMM to monthly data.

As previously noted, the parameter estimates and test statistics depend on the small sample properties of GMM; these results should be viewed with caution. The finite sample properties of GMM in the asset pricing framework are not clear if the true data generating process is like that of the Kandel/Stambaugh model.

5. MONTE CARLO RESULTS FROM THE MODEL ECONOMY

The object of this exercise was to examine the finite sample properties of GMM in the Kandel/Stambaugh model economy. In particular, to see what lessons we can draw about the properties of GMM in possibly similar environments in the real world.

5.1 Estimators

The performance of four GMM estimators on simulated data is examined. Estimators are distinguished by the asset returns and instrument set they use and by the number of iterations over the weighting matrix they permit. Table 3 describes the estimators used on simulated data. The first pair of estimators (TS1 and MS1) used only the riskless bond as an asset return and a constant, lagged consumption growth and the lagged return on the riskless bond as instruments. The second set of two estimators (TS2 and MS2) used the risk free bond, a share of aggregate wealth and a share of levered equity as the asset returns and a constant, lagged consumption growth and asset returns as the instruments. TS1 and TS2 used two-stage GMM. MS1 and MS2 used multi-stage GMM.

A maximum of 75 iterations over the weighting matrix was permitted for the multi-stage GMM estimators. This constraint was not often binding for data sets over 120 observations. Also, a maximum of 200 iterations to numerically minimize the quadratic (equation (16)) was permitted. For all estimations, the initial values for the parameters were the values of the parameters of the model economy, β=.99731, α= 55.

5.2 Small Sample Results

2000 samples of 120 observations were drawn from the Kandel and Stambaugh representative agent model economy. The models were estimated by the GMM with the four estimators described above. The results for this sample size are shown in Table 4.

The most prominent result confirms the finding of Kocherlakota in that there is a strong tendency for all estimators to "overreject" the model. Examining the last columns of Table 4, the actual rejection rates for the test of overidentifying restrictions are far higher in every case than the corresponding nominal size. For example, in the case of estimator MS2 the actual rejection rate is 43.7% at a nominal 1% size! The overrejection is even stronger than that found by Kocherlakota (1990) in a different environment. Rejection of the model is even more likely for the second two estimators (TS2 and MS2) which use more information.

The very strong overrejections of the model by GMM indicates that use of the asymptotic critical values in such an asset pricing environment may be inappropriate and the rejections of the representative agent model must be viewed with circumspection.

Two reasons for the strong overrejections will be considered, poor estimation of the parameters and the skewed and kurtotic distribution of the pricing errors.

Kocherlakota (1990) argued that the overrejection he found in his environment was due to poor estimation of the parameters. I find this is an even greater problem in the Kandel and Stambaugh environment than it was in the Mehra and Prescott environment. As found by Kocherlakota (1990), the estimates for α are strongly biased downwards in small samples. This suggests estimates of relative risk aversion based on this type of exercise will tend to understate risk aversion. The estimates for β are strongly biased upwards. The confidence intervals for both parameters are highly misleading. The extent of these problems is shown in the statistics in Table 4. The nominal 95% confidence intervals for α estimates cover the true value of α only 16-25% of the time. The coverage for the β confidence interval estimates are also poor, but not uniformly so, ranging from 30-91%. Figure 1 illustrates the frequencies of the estimates of the pairs of parameters for the

MS1 estimator for data sets of various lengths. The shape of the three dimensional histograms suggests a strong nonlinear relation between the parameter estimators.

Examination of the negative of the log of the simplified criterion function and its contour plot (shown in Figure 2) for the risk free asset and return to aggregate wealth confirms this suspicion. Figure 2 was constructed by numerical integration of the criterion function implied by the pricing errors of the riskless return and the return to aggregate wealth. The identity matrix was used as the weighting matrix. (Alternatively, similar figures could be constructed using very large samples of simulated data with any of the estimators used here.) Clearly, there is a very strong nonlinear relation between the GMM parameter estimates in this environment.

The "U-shaped ridge" in the criterion function is caused by the continuous distribution of consumption growth. To see this, consider the simple criterion function in which we treat β and the return on the asset as constants for tractability.

$$CF_t = \beta^2 \cdot (E_{t-1}(CG_t^{-\alpha}))^2 \cdot R_t^2 - 2 \cdot \beta \cdot (E_{t-1}(CG_t^{-\alpha})) \cdot R_t + 1 \tag{18}$$

Recalling that consumption growth is lognormally distributed, using the moment generating function of a normal distribution and taking a derivative with respect to α we get

$$\frac{\partial CF}{\partial \alpha} = 2\cdot\beta\cdot R_t\cdot(\beta\cdot R_t\cdot\exp(-\alpha\cdot\mu+.5\cdot(-\alpha)^2\cdot\sigma^2)-1)$$
$$\cdot\exp(-\alpha\cdot\mu + .5\cdot(-\alpha)^2\cdot\sigma^2)\cdot(-\mu+\alpha\cdot\sigma^2) \tag{19}$$

In general, the first term in the expression will have more than one value of α that sets it equal to zero and those values will depend on the value of β. For instance, assuming $\mu = .0049$, $\sigma = .0128$, $\beta = .99731$ and $R_t = 1.025$, $\alpha \approx 4.74$ and $\alpha = 55$ will set the first term on the right hand side of equation (19) to zero. The third term provides another turning point; it is set to zero if $\alpha = \mu/\sigma^2 \approx 29$.

If consumption growth were not lognormally distributed, if the variance were driven to zero - discrete valued as in the Mehra and Prescott environment studied by Tauchen (1986) and Kocherlakota (1990) - the expectation of consumption growth in equation (18) would just be a constant and the derivative with respect to α would be

$$\frac{\partial CF}{\partial \alpha} = 2\cdot\beta\cdot R_t(\beta\cdot R_t\cdot CG^{-\alpha} - 1)\cdot(-\ln CG)\cdot CG^{-\alpha} \tag{20}$$

For a given value of β, there is only one value of α which sets (20) equal to zero and hence there is no "U-shaped ridge" in the Mehra and Prescott criterion function. There is, however, a more conventional straight ridge in this case.

To combat estimation problems influencing the test of overidentifying restrictions, Kocherlakota (1990) suggests the Hansen-Jaganathan (1989) estimation-free method for testing whether overidentifying restrictions are met by a particular data set. Hansen and Jaganathan (1989) suggest treating each parameter specification as a different model and testing whether each specification of interest can satisfy the overidentifying restrictions. That is, a particular parameter set is chosen and the J statistic is constructed from the orthogonality conditions implied by the data and the chosen parameter values. Because the parameters are chosen a priori, the test statistic is distributed as a chi-square random variable with NK degrees of freedom.

To determine if estimation of the parameter values is, in fact, causing the overrejections, 1500 random sample of various sizes were drawn and the estimation-free J statistics were constructed with the true parameter values using all assets and instruments. If the estimation-free J statistics are truly chi-square NK, the order statistics from their p-values should lie on the forty-five degree line. If the p-values lie under the forty-five degree line, it indicates overrejection. If the p-values lie above the forty-five degree, it indicates underrejection. Figure 3 indicates that the J statistics converge relatively quickly to their asymptotic distribution, but that there is a fairly strong tendency to overreject. Estimation of the parameters does contribute to the severe overrejections in small samples. The second factor in the overrejections may be the distributions of the pricing errors (the orthogonality conditions implied by equation (3)) which are highly skewed and kurtotic. (See Kocherlakota (1993) for a discussion of kurtosis in tests of asset pricing models.) If the pricing errors are kurtotic, the central limit theorem on which the J statistic implicitly relies will not be a good approximation in small samples, there will be too many outliers and the statistic will tend to overreject.

In the simulated data, skewness and kurtosis are prominent features. For example, the coefficients of skewness and kurtosis constructed by numerical integration of the pricing error associated with the risk free asset or are approximately -3 and 21 respectively. In contrast, the coefficients of skewness and kurtosis for a normal distribution are 0 and 3 respectively. The distribution of the pricing error implied by the risk free asset is shown in Figure 4. Skewness and excess kurtosis are also prominent features of pricing errors in real data and so could be contributing to rejections in the real data.

5.3 Small to Large Sample Results

Table 5 shows the effect of increasing sample size on the GMM estimators. The true sizes of tests of overidentifying restrictions, the median point estimates and the confidence intervals for α (the coefficient of risk aversion) slowly converge to their asymptotic properties. Disturbingly, the point estimates of and confidence intervals for β remain poor even with very large samples. In fact, the estimates of β can actually become poorer as the sample size increases. Examination of the contour map of the criterion function in Figure 3 exposes the source of the problem, however. As the sample size increases and the median α estimates converge on their true value of 55, the β estimates actually move along the U-shaped ridge away from the true value of β.

5.4 Two-Stage Versus Multi-Stage GMM

The true values of the parameters are used as the starting values for the estimators in the Monte Carlo experiments to give the estimators the benefit of the doubt. For both estimators and all samples sizes, the two-stage procedure has better properties in terms of the tests of overidentifying restrictions and parameter estimates than the multi-stage procedure. It is not clear whether this is a result of using the true parameter values as the starting values. The two-stage procedure may reduce the estimation problems in a manner similar to that of the estimation-free Hansen and Jaganathan method. Further work with a variety of starting values may resolve this question.

5.5 First Versus Second Estimators

The two estimators used in the Monte Carlo experiments were chosen to mimic a cross section of the estimators that provided strong rejections in the real data. The first estimators (TS1 and MS1) - with only the riskless return and the smaller instrument set - rejected less often in small samples, in contrast to real data. The second set of estimators (TS2 and MS2) which had all three asset returns and a larger instrument set, dominated the first set both with respect to correct size and coverage of the confidence intervals around parameter estimates as sample size increased. The two-stage procedure with the second estimator (TS2) has the best GMM performance in terms of rejection rates and estimates of the parameters.

These results suggest that in small samples, more orthogonality conditions may hurt the performance of GMM in this environment and lead to erroneous inference.

6. CONCLUSIONS

Kandel and Stambaugh (1990) and Cecchetti, Lam and Mark (1990) have created models of the financial economy which illustrate some valuable lessons on the type of data consistent with rational asset pricing models. This paper has investigated the properties of GMM in such an environment. It is found that GMM procedures strongly overreject the model and provide poor point estimates. The strong tendency to overreject the true model means that we should view the rejections of such models with real data with circumspection.

The tendency to overreject in this environment is partially explained by poor parameter estimation which is caused by a U-shaped ridge in the criterion function. Due to the unusual U-shaped ridge, the GMM point estimates for α were found to be significantly biased downwards, consistent with Kocherlakota's results. Such bias indicates that there is merit to Kandel and Stambaugh's claim that the true value of the coefficient of relative risk aversion may be substantially higher than commonly thought. In addition, due to the peculiar shape of the criterion function, the confidence intervals for β (the discount factor) remain misleading even at very large sample sizes.

Poor parameter estimation is not the whole story behind the overrejections of the model as examination of the Hansen and Jaganathan estimation-free J statistics show. Even if the value of the parameters is specified to be the true value a priori, the J statistic still tends to overreject the model in small samples.

The overrejections of the model may also be partially caused by the skewness and kurtosis of the pricing error data. This feature of the simulated data is shared by data from the real world.

Two-stage estimators were found to be superior to multi-stage estimators in terms of rejection rates and coverage of parameter estimate confidence intervals. As the GMM algorithm was started at the true parameter values, it is not clear if this superiority is simply an artifact of the parameter estimation problems.

In small samples, the first estimator with one asset return and a smaller instrument set performed better than the second estimator. This advantage was reversed as sample size increased, however, leading to no clear conclusions for moderate sized samples.

APPENDIX: THE KANDEL AND STAMBAUGH MODEL ECONOMY

Table A.1 - Markov Transition Probability Matrix for the State of Nature

State	1	2	3	4
1	.845	.113	.042	.000
2	.107	.750	.054	.089
3	.067	.067	.800	.067
4	.026	.079	.105	.789

Table A.2 - Parameters of the Distribution of Consumption Growth for Each State of Nature

State	μ	σ
1	.0049	.0128
2	.0041	.0128
3	.0049	.0131
4	.0041	.0131

REFERENCES

Cecchetti, S. G., Lam, P., and Mark, N. C. (1990), "Mean Reversion in Equilibrium Asset Prices," *American Economic Review*, 3, 398-418.

Hansen, L. (1982), "Large Sample Properties of Generalized Method of Moment Models," Econometrica, 50, 1029-1053.

Hansen, L., and Jaganathan, R. (1989), "Restrictions on Intertemporal Marginal Rates of Substitution Implied by Asset Returns," Working Paper 12, Northwestern University, Department of Finance.

Hansen, L., and Singleton, K. (1982), "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269-1286.

_____ (1984), Errata, *Econometrica,* 52, 267-268.

Kandel, S., and Stambaugh, R. F. (1988), "A Model of Expected Returns and Volatilities for Various Forecast Horizons," unpublished manuscript.

_____ (1990), "Expectations and Volatility of Consumption and Asset Returns," *The Review of Financial Studies*, 3, 207-232.

Kocherlakota, N. R. (1990), "On Tests of Representative Consumer Asset Pricing Models," *Journal of Monetary Economics*, 26, 285-304.

_____ (1993), "Testing the Consumption CAPM with Heavy Tailed Pricing Errors," *University of Iowa Working Paper*, 93-10.

Lucas, R. (1978), "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429-1455.

Mehra, R., and Prescott, E. (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145-161.

Tauchen, G. (1986), "Statistical Properties of Generalized Method of Moment Models of Structural Parameters Obtained from Financial Market Data," *Journal of Business and Economic Statistics*, 4, 397-425.

Table 1: Model descriptions for GMM estimation on real data

Estimator	Asset Returns	Instrument Set	Maximum Iterations Over W
TR1	T-bill	constant, CG_{-1} & T-bill_{-1}	2
MR1	T-bill	constant, CG_{-1} & T-bill_{-1}	75
TR2	VWNYSE	constant, CG_{-1} & VWNYSE_{-1}	2
MR2	VWNYSE	constant, CG_{-1} & VWNYSE_{-1}	75
TR3	T-bill & VWNYSE	constant, CG_{-1},T-bill_{-1} & VWNYSE_{-1}	2
MR3	T-bill & VWNYSE	constant, CG_{-1},T-bill_{-1} & VWNYSE_{-1}	75

Table 2: Results from GMM estimation on real data

Estimator	sample size	d.f.	Alpha (s.e.)	Beta (s.e.)	J statistic (p-value)
TR1	120	1	-2.95 (0.71)	0.97 (0.00)	7.98 (0.00)
MR1	120	1	-2.94 (0.66)	0.97 (0.00)	8.70 (0.00)
TR2	120	1	-4.35 (5.49)	0.95 (0.03)	2.89 (0.09)
MR2	120	1	-4.58 (4.85)	0.95 (0.03)	2.33 (0.13)
TR3	120	6	1.23 (20.87)	0.97 (0.12)	5.99 (0.42)
MR3	120	6	-2.68 (0.57)	0.97 (0.00)	16.12 (0.01)

Table 3: Model descriptions for GMM estimation on simulated data

Estimator	Asset Returns	Instrument Set	Maximum Iterations Over W
TS1	Risk Free	constant, CG_{-1} & Risk $Free_{-1}$	2
MS1	Risk Free	constant, CG_{-1} & Risk $Free_{-1}$	75
TS2	Risk Free, Aggregate Wealth and Levered Equity	constant, CG_{-1}, Risk $Free_{-1}$ Aggregate $Wealth_{-1}$ & Levered $Equity_{-1}$	2
MS2	Risk Free, Aggregate Wealth and Levered Equity	constant, CG_{-1}, Risk $Free_{-1}$ Aggregate $Wealth_{-1}$ & Levered $Equity_{-1}$	75

Table 4: Results from GMM estimation on simulated data, small sample performance

sample size	Estimator	DF	Median Alpha	% of 95% C.I.'s covering true alpha	Median Beta	% of 95% C.I.'s covering true beta	Monte Carlo 10, 5, 1% critical values			Model Rejection Rates at Nominal Sizes of 10, 5, 1%		
120	TS1	1	3.51	22	1.01	91	6.0	13.2	39.0	20.7	15.4	9.7
120	MS1	1	3.55	17	1.01	84	21.0	30.5	53.0	38.8	33.5	25.8
120	TS2	13	33.14	25	1.03	55	31.5	41.3	68.8	34.3	24.4	13.5
120	MS2	13	9.28	16	1.01	30	54.1	69.3	82.6	67.0	56.9	43.7

Table 5: Results from GMM estimation on simulated data, small to large sample performance

sample size	Estimator	DF	Median Alpha	% of 95% C.I.'s covering true alpha	Median Beta	% of 95% C.I.'s covering true beta	Model Rejection Rates at Nominal Sizes of 10, 5, 1%		
480	TS1	1	11.01	35	1.04	68	17.9	13.0	8.2
1920	TS1	1	24.60	59	1.05	55	17.6	12.8	8.3
7680	TS1	1	40.69	76	1.04	62	14.6	9.6	4.7
480	MS1	1	11.05	27	1.04	54	35.4	30.6	23.0
1920	MS1	1	24.56	50	1.05	49	31.43	26.9	20.2
7680	MS1	1	40.57	71	1.04	58	23.5	19.3	12.3
480	TS2	13	41.78	49	1.04	60	21.2	13.1	4.7
1920	TS2	13	45.67	73	1.03	64	8.4	4.4	1.1
7680	TS2	13	51.78	89	1.01	81	8.9	4.7	2.3
480	MS2	13	16.46	20	1.02	25	61.7	54.7	44.5
1920	MS2	13	41.99	52	1.04	45	29.4	22.1	12.3
7680	MS2	13	51.66	85	1.01	78	14.7	9.1	5.8

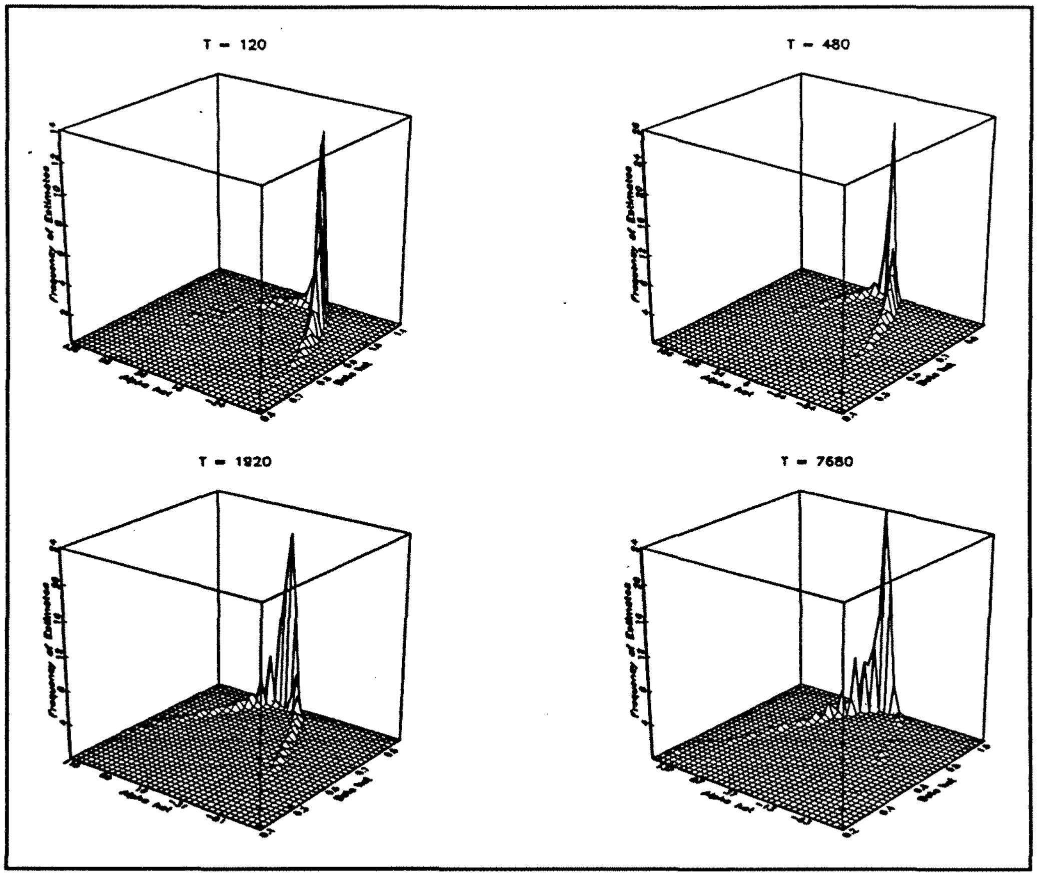

Figure 1
The frequencies of the parameter estimates from estimator MS1 on simulated data for various sample lengths. Note the strong nonlinear relationship between the parameter estimates.

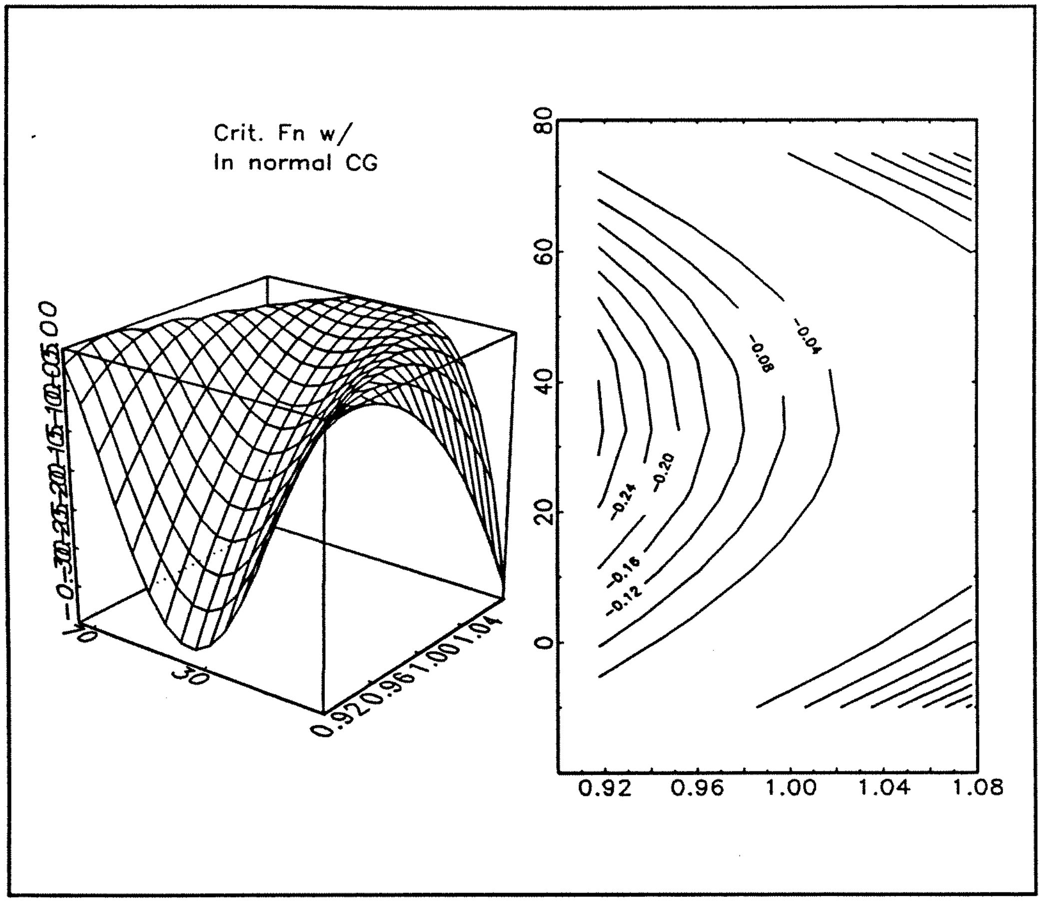

Figure 2
The log of the negative criterion function and its contour plot shows the "U-shaped" ridge due to the strong nonlinear relation between the parameter estimates. The figure was constructed by numerical integration of the pricing errors implied by the risk free bond and the return to aggregate wealth.

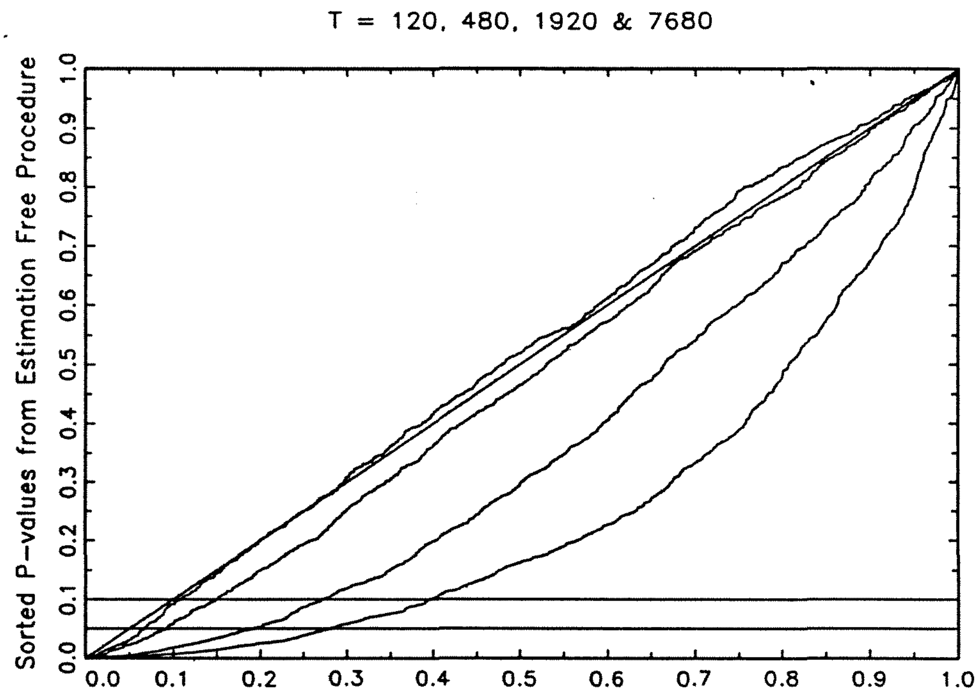

Figure 3

Sorted p-values from the "true" J statistics created from 1000 simulated data sets of various lengths. The J statistics are constructed with the "true" parameter values of the model economy. If the statistic is truly chi square with NK degrees of freedom, the sorted p-values should lie along the 45 degree line. The horizontal lines at .05 and .1 may be used to find the degree of overrejection for each sample size. For example, the actual rejection rate for a sample size of 120 observations at a 10% nominal size is approximately 40%. The strong overrejections indicate that poor parameter estimation is not the only contributing factor to the overrejection of the model, the kurtosis of the pricing errors also contributes.

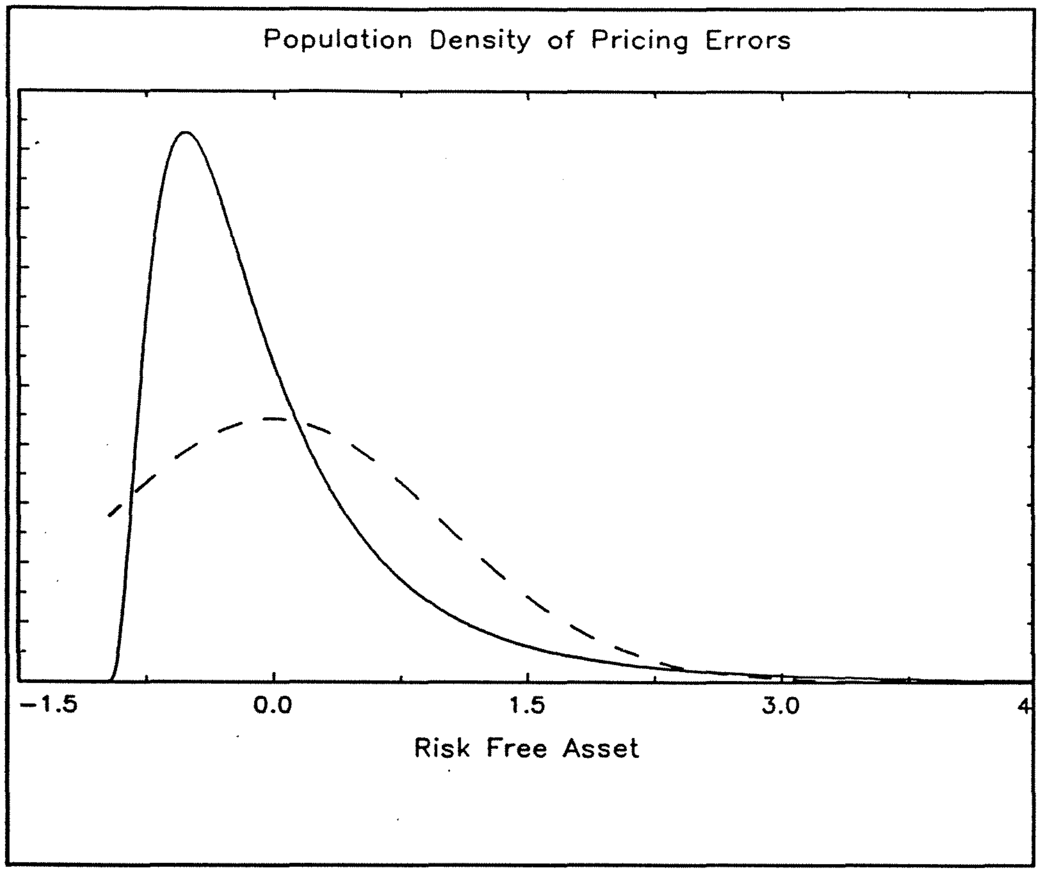

Figure 4
The population density of the pricing errors of the risk free asset overlaid by a normal density of the same mean and variance to highlight the skewness and kurtosis of the pricing error distribution. The distribution was constructed by numerical integration of the pricing error of the riskless bond.